Naval Aerospace Medical Research Laboratory

NAMRL-1401

THE FITNESS IMPAIRMENT TEST (FIT): A FIRST LOOK

C. A. Portman-Tiller

Naval Aerospace Medical Research Laboratory
51 Hovey Road
Pensacola, Florida 32508-1046

Reviewed and approved 27 Apr 98

L. H. FRANK, CAPT, MSC USN
Commanding Officer

This research was sponsored by the Naval Medical Research and Development Command under work unit 63706N M0096.002-7010 DN 248502.

The views expressed in this article are those of the authors and do not reflect the official policy or position of the Department of the Navy, Department of Defense, nor the U.S. Government.

Volunteer subjects were recruited, evaluated, and employed in accordance with the procedures specified in the Department of Defense Directive 3216.2 and Secretary of the Navy Instruction 3900.39 series. These instructions are based upon voluntary informed consent and meet or exceed the provisions of prevailing national and international guidelines.

Trade names of materials and/or products of commercial or nongovernment organizations are cited as needed for precision. These citations do not constitute official endorsement or approval of the use of such commercial materials and/or products.

NAVAL AEROSPACE MEDICAL RESEARCH LABORATORY
51 HOVEY ROAD, PENSACOLA, FL 32508-1046

NAMRL-1401

THE FITNESS IMPAIRMENT TEST (FIT): A FIRST LOOK

C. A. Portman-Tiller

ABSTRACT

In a technologically advanced society with increasing demands, shiftwork has become the norm. Although shiftwork has become an accepted way of life, it has disadvantages, such as increases in fatigue, depressed mood states, and decreases in performance. The object of the study was to validate a test developed to determine whether or not a worker is fit for duty. The Fitness Impairment Test (FIT) is based on ocular measures that are believed to be correlated with fatigue. Participants were 9 Navy and Marine Corps aviation candidates, ages 23 to 26 years. ANOVA and correlational analyses were used to compare subjective and objective measures of performance and fatigue with the measures produced by FIT. Although participants showed decreases in performance and mood states, few statistically significant correlations were found between the physiological and performance measures.

ACKNOWLEDGMENTS

Special thanks go out to the Marine Aviation Support Group (MATSG) at NAS Pensacola. Without their support, this research would have not been possible. The author would also like to thank LT Sean Biggerstaff, USN, MSC, Scott Meyer, Ph.D., and K. Mayer for their editorial contributions, and Drs. Doug Weigmann and Bob Stanny for their statistical guidance.

INTRODUCTION

In today's fast-paced, technological society, it has become critical for some organizations to maintain 24-hour schedules. These organizations include health care facilities, manufacturing companies, the military, the transportation industry, and the like. Furthermore, the increase in the demand for goods and services offered by these and other organizations, as well as, push from management to become more efficient, have contributed to the prevalence of shiftwork (Moore-Ede, 1993). Shiftwork refers to a number of work schedules in which the total time of production of commodities or services exceeds the number of hours in a normal 8-hour work day and engages more than one team or shift of workers (Akerstedt & Gillberg, 1982). Throughout this paper, shiftworkers refers to individuals who work rotating schedules, rather than a fixed schedule. For example, during one work week an individual may be on the day shift and the following week he/she may be switched to the night shift.

Rotating shifts create disruptions in the circadian rhythms that maintain a worker's wake/sleep/activity patterns. One result of disrupted circadian rhythms is loss of sleep, which leads to fatigue, which turn may lead to problems in performance, mood, health, and (Akerstedt, 1991; Krueger, 1989; Monk, 1991; Monk & Folkard, 1983; Moore-Ede, 1993; Nicholson & Stone, 1987). Figure 1 is flowchart demonstrating the relationship described above.

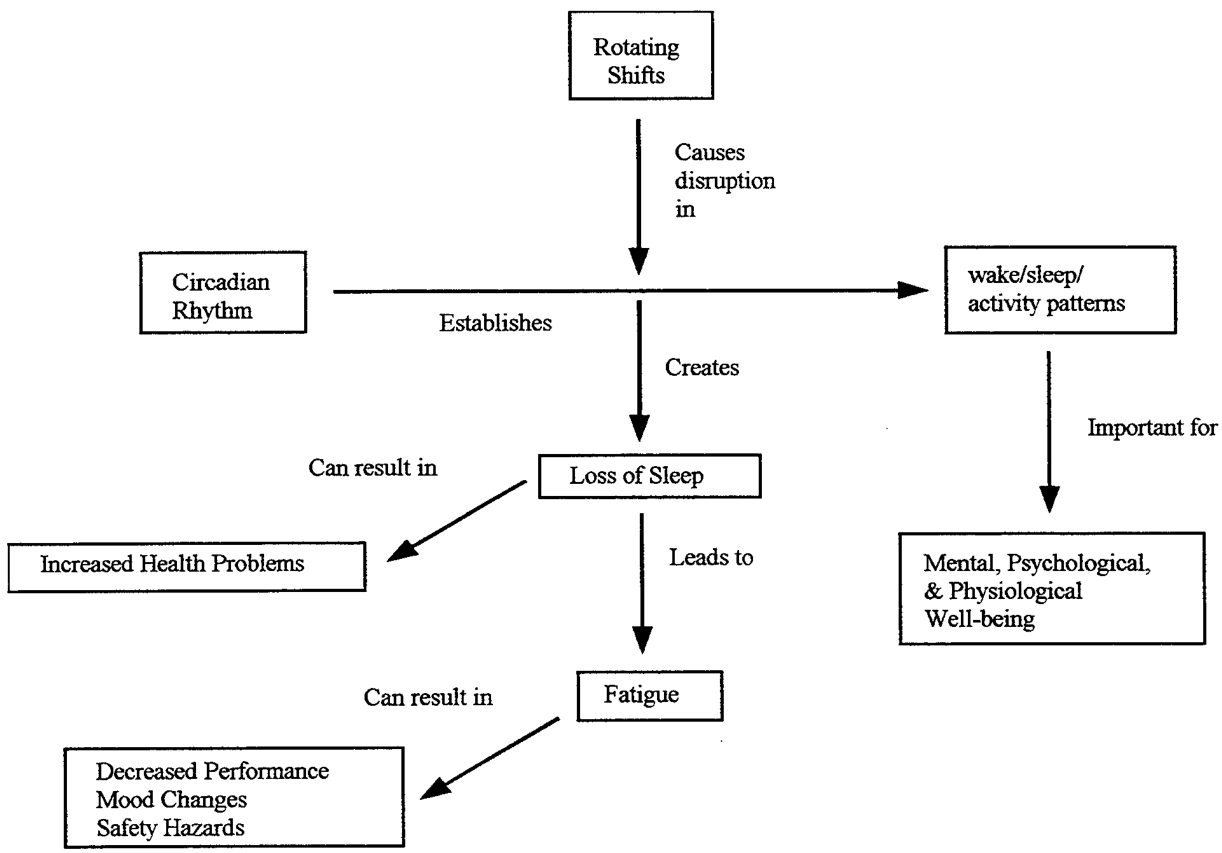

Figure 1. Relationship between circadian rhythms and rotating shifts.

Effects of Shiftwork

Shiftwork and circadian rhythms. One of the demonstrated effects of shiftwork is a disruption in a worker's circadian rhythms, "the non-trivial fluctuations in almost every physiological measure over a 24-hour period" (Monk & Folkard, 1983, p. 97). Circadian rhythms are free running, self sustaining fluctuations that are regulated by environmental cues such as light and wake/sleep schedules. If an individual does not have a set routine, as occurs when working rotating shifts, his/her rhythm may be disrupted (Nicholson & Stone, 1987). Because circadian rhythms establish a wake/sleep/activity pattern, disruption of these rhythms can have effects on an individual's psychological and physiological well-being. Research has shown that negative side-effects of a circadian rhythm disruption include loss of sleep, increased health problems, and depressed mood states (Monk, 1991; Monk & Folkard, 1983; Moore-Ede, 1993).

Shiftwork, sleep loss and health. People who work rotating shifts experience impairments in the quality and quantity of sleep (Akerstedt, 1991). When workers transition from day to night shifts, they frequently do not sleep during the day of the first night shift. This sleep loss can create a variety of problems, including reduced alertness, impaired performance, and lapses of consciousness (Akerstedt, 1991; Krueger, 1989; Moore-Ede, 1993).

According to Colligan and Tepas (1986), increased medical problems and costs are also associated with sleep loss attributed to the disruption of circadian rhythms. Individuals deprived of adequate sleep report higher rates of gastrointestinal and digestive disorders. Furthermore, there is the potential to exacerbate pre-existing medical illnesses (Colligan & Tepas, 1986). The dose-response functions for many prescription medications follow circadian rhythms, and prescribing physicians almost always assume that a patient is on a diurnal (daytime) schedule (Colligan & Tepas, 1986). Therefore, medications designed to work with what would be the body's natural rhythms may sometimes do more harm than good (Colligan & Tepas, 1986).

People who experience sleep deprivation frequently fail to attend to critical tasks, and their powers of reasoning and creativity may be diminished (Moore-Ede, 1993). Akerstedt (1991) estimated that 75% of all night workers experience fatigue, and that 20% of these actually fall asleep on the job. A review of the literature by Krueger (1989) found that "sleep loss appears to result in reduced reaction time, decreased vigilance, perceptual and cognitive distortions, and changes in affect" (p. 129).

Fatigue, performance, and mood. Numerous researchers have investigated the effects of fatigue on performance. For example, Mast and Heimstra (1964) examined the impact of fatigue on vigilance performance. Vigilance is defined as alert watchfulness. Mast and Heimstra found a large decrement in vigilance performance for participants in the mental fatigue condition and a small decrement for participants in the non-fatigue condition.

In a related study, Angus, Heslegrave, and Myles (1985) explored the effects of prolonged sleep deprivation on mood and performance in participants with and without exercise. Participants were deprived of sleep for a 60-hr period on two separate occasions. Angus et al. found decrements in mood and performance resulting from the sleep loss. These decrements were found to be independent of exercise. Exercise did not delay or offset the effects of fatigue. Studies at the Naval Aerospace Medical Research Laboratory, Pensacola, FL, have also shown that performance and mood decrease as a result of sleep deprivation (Stanny, McCardie, & Neri, 1993; Wiegmann et al., 1993). In these studies, participants reported increased fatigue, boredom, sleepiness, and decrease in attention and alertness across time.

Shiftwork and safety. Safety of employees has also become an important issue for researchers and corporations alike. In the field of aviation, human error is now believed to cause 66% of all commercial airline accidents, 79% of commuter accidents, and 88% of all private accidents. These mishaps have been attributed to both pilot error and air traffic controller error, two professions in which shiftwork and circadian rhythm disruptions are common (Moore-Ede, 1993). Fatigue has been suspected in many other real world accidents, including the Exxon Valdez oil spill and the nuclear reactor failure at Three Mile Island (Rosekind et al., 1994).

Gold et al., (1992) found that nurses working shifts reported more disruption of their sleep/wake cycles and reported dozing off more at work than nurses who worked steady shifts. The researchers also found that rotating shift nurses were twice as likely to fall asleep while driving to and from work, and were twice as likely to report accidents or errors related to sleepiness.

Fitness-for-Duty Tests

The possibility of human-error related accidents and the clear potential for additional mishaps makes study of the effects of fatigue of major practical concern for industry and the military. The need to operate on a 24-hour schedule has not and will not decrease for either group in the near future. Therefore, it is necessary for some employers to develop tests to determine whether a person is fit to work their shift to ensure the safety and well-being of the employee, employer, and the consumer. Workers might be tested before and after work to ensure their safety on the job and in transit. Moreover, people could be tested during work to detect impairment during their shifts. Devices that measure fatigue-related performance decrements are called Fitness-for-Duty or Readiness-to-Perform tests; a number of such tests are commercially available (Gilliland & Schlegel, 1994; Moore-Ede, 1993).

Performance based tests. One Fitness-for-Duty test is called CogScreen (Kay, 1995). This is a computerized test that measures accuracy and reaction times on a series of cognitive tests. CogScreen was originally developed by the Federal Aviation Administration (FAA) and is currently used in the "medical re-certification evaluation of pilots with known or suspected neurological and/or psychiatric conditions" (Kay, 1995, p. 6). Because it takes approximately 45 min to take CogScreen, it is administered only when there is a suspected problem. When impairment is not obvious to others, individuals are not tested; consequently, CogScreen will not reveal unsuspected impairments.

Another performance-based test is FACTOR 1000 (Performance Factors, Inc., 1993). Factor 1000 is a video tracking test that measures eye-hand coordination, visual perception, and reaction time. This test takes less than one minute to administer. Employees' scores are compared to their own baselines established in previous testing. Validity coefficients for the test are not explicitly stated; however, the tests developer refers to validation of the task with alcohol-induced performance decrements. The test detected significant decreases in scores at blood alcohol concentrations (BAC) below 0.05% in about half of the validation sample. When the BAC was greater than 0.10%, over 90% of the participants had decrements in performance on the Factor 1000 test (Performance Factors, Inc, 1993).

Unfortunately, there are potential problems with most performance-based Fitness-for-Duty tests. First, test administration times may be considerable. Workers, of course, could be tested during normal working hours, decreasing the productive man-hours for the organization. Shorter tests, however, may be less accurate because an impaired employee may be able to sustain performance long enough to pass a brief test. Third, performance is evaluated against a baseline that the individual has set himself. Workers might intentionally set a low baseline so that the test will not detect impairment. Finally, many Fitness-for-Duty tests were developed and validated for the purposes of detecting drug use and/or impairment due to drug use and not impairment due to fatigue - with the exception of CogScreen, which was developed to detect neurological impairment (Kay, 1995). The criterion used to detect drug impairment is different from the criterion that is used to detect fatigue impairment. Drug testing, whether it be alcohol, cocaine, or some other drug, is normally done by blood draws or saliva samples. It is not possible, as of yet, to test for impairment due to fatigue using a chemically-based test. Therefore, some other way of detecting impairment due to fatigue needs to be investigated.

Physiologically based test. A physiologically based test may overcome some of these problems due to the involuntary nature of many physiological responses. By focusing on involuntary responses, a test might more accurately measure an individual's level of impairment. The most frequently used types of involuntary responses are pupillary responses, and galvanic skin responses, as used with lie detectors.

Research on the relationship between fatigue and pupillary responses has shown that in a well rested state the pupil is at its maximal diameter. However, pupil diameter decreases with fatigue and reaches minimal diameter just before the onset of sleep (Andreassi, 1989). Thus, a test based on pupillary responses might be a reliable measure of impairment caused by fatigue.

For decades, pupillary responses have been used as measures of performance and emotion (Andreassi, 1989). An example of an early method used to measure pupil size is infrared photography, which allowed the experimenter to take a picture of the pupil and its diameter. A photoelectric method was also developed, which measured the light reflected from the iris to determine the pupil size (Andreassi, 1989). These methods were used for years, despite the fact that they were cumbersome and not very accurate. The Lowenstein pupillograph, which uses infrared scanning of the iris to

determine the amount of light reflected, eventually replaced these methods. The pupillograph measures pupil diameter and also gives information about the rate of change in the pupil size. Although the Lowenstein pupillograph was an improvement over the other two devices, accuracy has been a problem (Andreassi, 1989).

The video-based pupillometer is the most recent apparatus developed that takes into consideration the need for practicality and reliability, and is currently the most widely used of the measurement devices. The video-based pupillometer is an electronic system that uses a closed circuit TV system to observe the eye, and a signal processor to measure and display the pupil diameter (Andreassi, 1989).

Only within the last 30 years have researchers had the ability to develop physiological devices that are both practical and reliable (Andreassi, 1989). The Fitness Impairment Testing (FIT) is such a device. FIT is a 30-second test that monitors changes in the central nervous system. Various ocular parameters such as pupil diameter and saccadic velocity, are used by the FIT to detect cognitive impairments such as fatigue (Pulse Medical Instruments, 1994). Because the reactions to the stimuli presented are involuntary, responses are presumably unaltered by learning, motivation, or any of the problems discussed previously.

METHODS

Participants

Volunteer participants were 9 Naval and Marine Corps aviation candidates, 23 to 26 years of age ($\underline{M}$ = 24.11, $\underline{SD}$ = 0.99), who were awaiting the start of initial flight training at Pensacola Naval Air Station. Participants were medically examined prior to and after the experiment. All participants were completely informed as to the purpose of the study, and of their freedom to withdraw at any time without prejudice to their military careers.

Materials

Fitness Impairment Test (FIT). The FIT is a fitness-for-duty test that measures changes in pupillary responses and eye movements. The changes measured indicate increases or decreases in pupil dilation or slowing in saccadic movements, and are identified by examining the reactions of the eye to light stimuli. These responses are believed to be related to changes in the central nervous system. Participants look through an eye piece and move their eye to follow a small point of light that moves from left to right. Brief flashes of moderate intensity light are then presented. The device records movements of the participant's eye and the dilation and constriction of the pupil.

Visual Analog Scale (VAS). The Visual Analog Scale is a customized symptoms inventory that assesses mood states and physical symptoms (DeJohn, Marr, Molina, & McCardie, 1992). Some of the symptoms included in the scale are depression, boredom, sleepiness, fatigue, hunger, and thirst. A sequence of these symptoms is displayed above a horizontal line segment. The left end of the line is labeled with a "0", the right end is labeled with a "100." Participants are instructed that the "0" means that they are not experiencing the symptom and "100" means that they are experiencing the most intense sensation they could imagine. An arrow appears in the middle of the line segment and participants are asked to use a trackball to move the arrow to the position on the line segment that corresponds to the intensity of the symptom they are experiencing at that moment. It takes approximately 4 min to take the VAS.

Running memory task. The running memory task measures vigilance and attention. This task requires participants to monitor a randomly ordered sequence of upper case letters presented individually for 50 ms each in the center of the video display. The interstimulus interval is 1250 ms, measured from stimulus onset to stimulus onset. Participants are required to decide whether each letter matches the immediately preceding letter in the sequence. Letters are presented randomly with a 0.50 probability that any letter presented will match the previous letter. Participants respond by pressing a key with the first finger of one hand whenever a letter matches its immediate predecessor. Participants press another key with the first finger of the opposite hand whenever a letter does not match the immediate predecessor. Key assignments to dominant and nondominant hands were counterbalanced across participants. A total of 480 stimuli were presented in each block of trials, requiring approximately 10 min to complete.

Because this study was a component of a larger research project, participants were also exposed to other tasks, including a memory search task, unstable tracking task, and several physiological measurements (e.g., body temperature). However, the focus of the current research was limited to the measures described above.

The running memory task was chosen instead of the unstable tracking and the memory search tasks because past research has indicated that the measures recorded from the running memory task (lapses, d', reaction time) positively correlate with the measures recorded from the unstable tracking and memory search combined (Stanny et al., 1993; Wiegmann et al., 1993). Therefore, instead of using two tests, one can be used and the same conclusions can be drawn. The measure of d' takes into account each participant's correct responses and error rates. It is a commonly used measure based upon signal detection theory.

Relationship between measures. A running memory task was used as a performance measure to determine criterion validity. The reason the VAS was chosen is that some components of the VAS could be used to establish convergent validity, and other components could be used to establish divergent validity.

Procedures

Training phase. Participants reported on a Monday morning at 0800 for a briefing in which they were informed as to the purpose of the study and that their participation was voluntary. Participants signed the consent forms and were given a medical examination. At 0900 participants began their first practice session. For 4 consecutive days, participants completed two practice blocks per day, each lasting approximately 90 min, including a 20-min break (see Table 1). In addition, there were three 3-min breaks included in each block.

Table 1. Experimental Task Order and Duration 1 of 9 Repeated Blocks

Task	**Duration** (min)
VAS	4
Tracking	10
Break	3
Running Memory	10
Break	3
Memory Search	16
Break	3
Memory Search	16
FIT	3
Break	20

Testing phase. On the last day of training (Thursday), participants were released at their normal time and asked to return at 1930 to begin the overnight testing. Participants were also asked not to nap before the evening testing session in order to simulate a normal day's work. Conditions during the experiment were identical to those during the training phase. However, testing continued until approximately 0900 Friday morning, by which time participants had completed nine blocks of testing. Participants were provided with caffeine-free food and beverages during their 20-min breaks. By the end of the testing phase, participants had been awake for at least 24 hrs.

After completion of the final block, participants were asked to stay in the dormitory to sleep for approximately 6 hours. After participants rested, they were debriefed and checked again by the medical monitor. Participants were then released.

Scoring. The index obtained from the FIT was correlated with the running memory task in an attempt to determine whether performance had declined. A global FIT index[1] was calculated as follows:

Index = (INI - mnpracINI)/stdpracINI) + (AMP - mnpracAMP)/stdpracAMP) +
(LAT - mnpracLAT)/stdpracLAT) + (SV - mnpracSV)/stdpracSV).

In this equation mnprac is the mean of all the practice sessions, and stdprac is the standard deviation for all the practice sessions. INI is the initial diameter of the pupil, measured just prior to the light flash, AMP is the constriction amplitude of the pupil defined as the difference between the initial diameter and the minimum diameter reached after the flash. (This value is averaged over the four flashes.) LAT is the latency of the pupil constriction or the time between the beginning of the flash and the beginning of the pupil constriction, and SV is the saccadic velocity (how fast the eye moved from side to side) (Pulse Medical Instruments, 1994).

Design

This was a one-way repeated measures experiment in which the independent variable was time of day during which participants performed the tests. There were nine blocks, in which a variety of variables were measured at varying levels of fatigue.

A correlational analysis was also performed using the FIT index and the VAS measures. There are a total of 28 VAS symptoms (see Appendix A for a list of all symptoms). Only eight were chosen as correlates of fatigue. Past research (Wiegmann et al., 1993) indicates that fatigue, boredom, depression, and sleepiness were positively correlated with sustained wakefulness or fatigue (providing convergent validity) and that attention, alertness, euphoria, and talkativeness were negatively correlated with sustained wakefulness or fatigue (providing divergent validity).

RESULTS

The first stage of the analysis focused on establishing that fatigue had been successfully manipulated. Subsequent stages address specific questions about the pattern of correlations among physiological and performance measures. The manipulation check was done by evaluating three measures - the performance measure, the subjective mood scale and the sleepiness scale using one-way ANOVAs. Throughout the paper, results reported as statistically reliable achieved significance levels at or beyond an alpha level of .01.

Results of the running memory task indicate that time had reliable effects on accuracy (d'), $F(1, 4.27) = 23.06$, $MSE = 5.20$, RT, $F(1, 7.01) = 11.12$, $MSE = 23703.34$, and the number of lapses, $F(1, 2.58) = 12.92$, $MSE = .30$ (see Figures 2 through 4)[2].

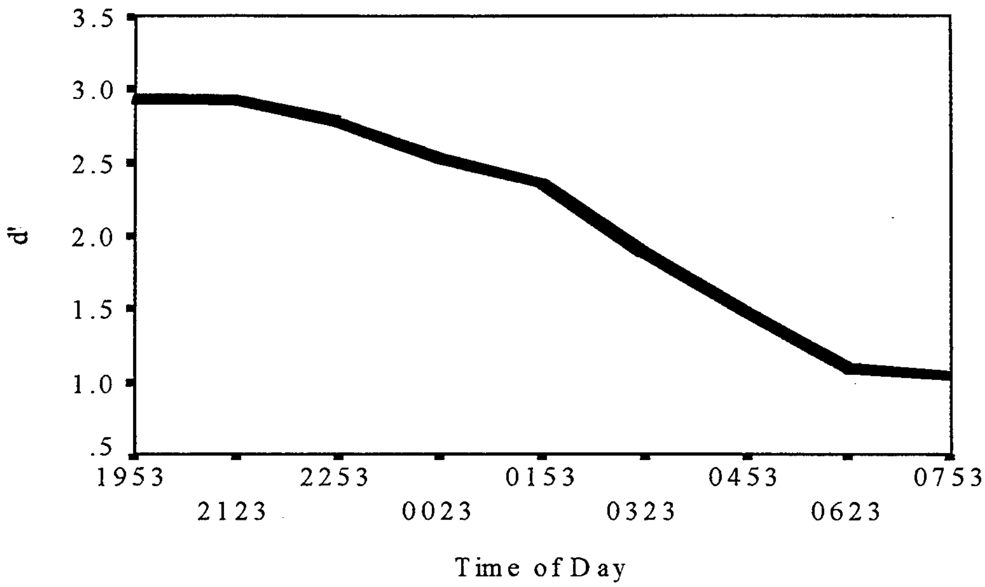

Figure 2. The measure of d' for the running memory task as a function of time.

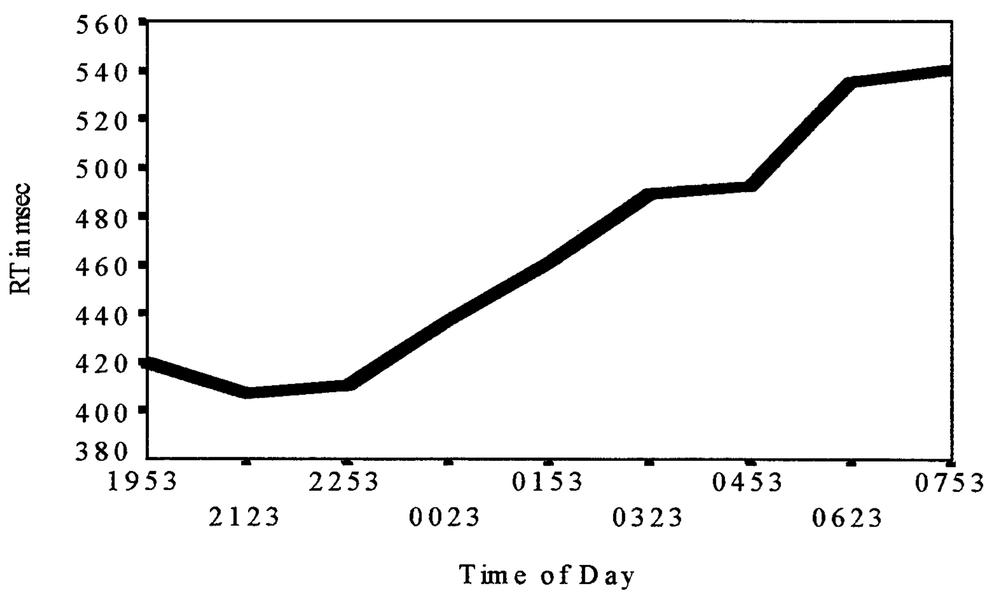

Figure 3. Reaction time for the running memory task as a function of time.

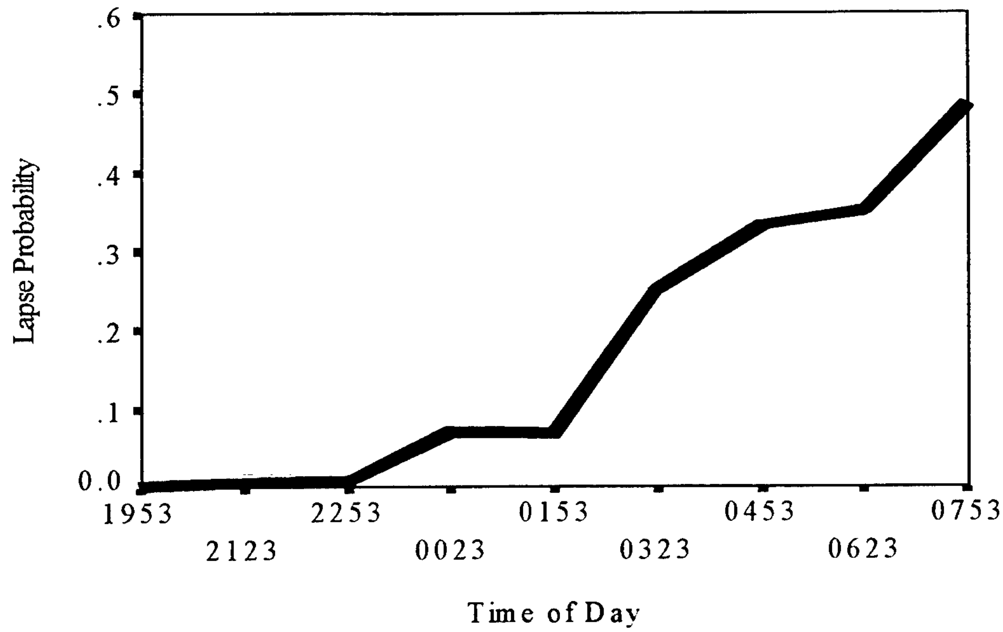

Figure 4. Lapse probabilities for the running memory task as a function of time.

The VAS also showed reliable changes for five of the eight symptoms chosen for review. The following five symptoms produced reliable changes: attention, alertness, fatigue, sleepiness, and boredom. Symptoms that decreased in magnitude over blocks are presented graphically in Figure 5; symptoms that increased in magnitude are presented in Figure 6. The self-reported measure of attention decreased across time ($\underline{F}$ (1, 3.55) = 4.71, MSE = 2806.97) and alertness also decreased as a function of time ($\underline{F}$ (1, 1.74) = 9.11, MSE = 3131.63).

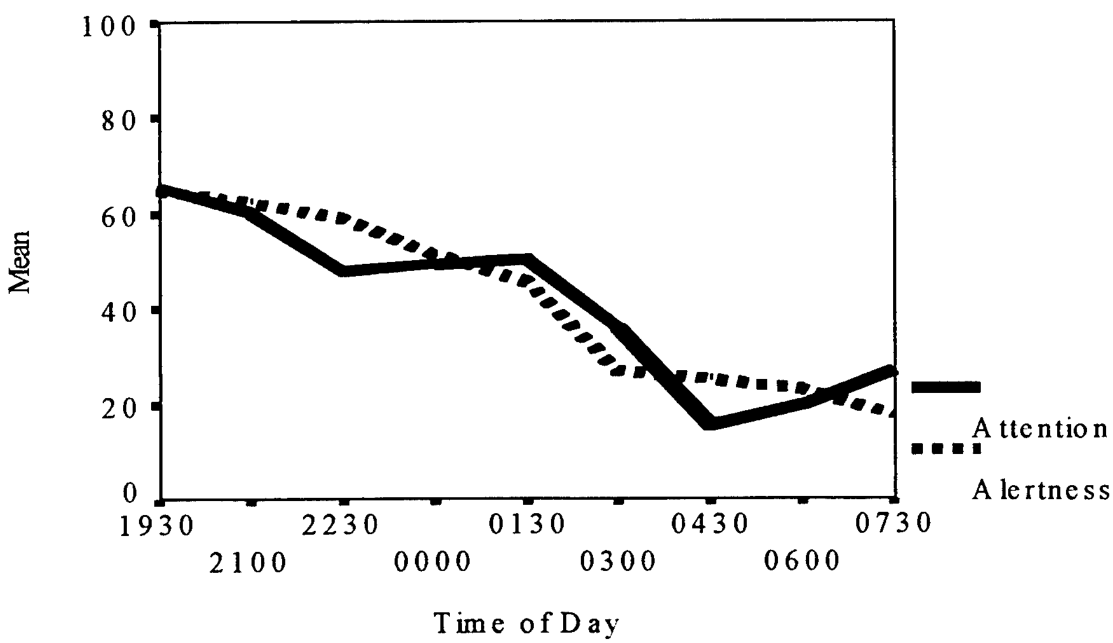

Figure 5. VAS symptoms that significantly decreased over time.

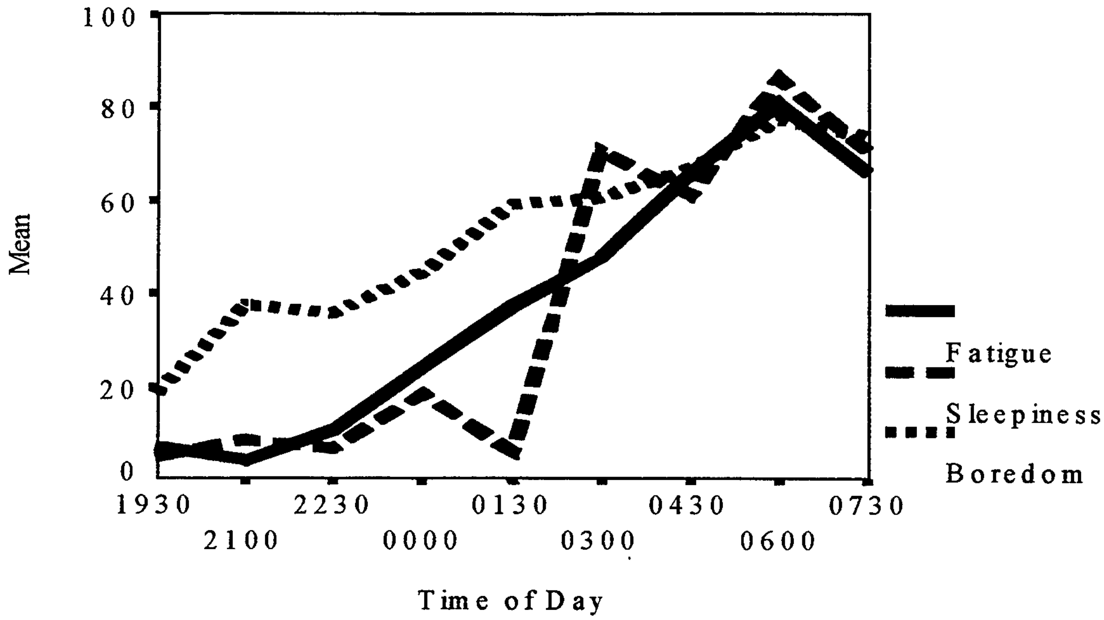

Figure 6. VAS symptoms that significantly increased over time.

Similarly, self-reports of fatigue (F (1, 3.34) = 13.77, MSE = 7382.08) and sleepiness increased across time (F (1, 6.68) = 18.64, MSE = 9334.23). Overall, participants reported significantly less sleepiness in blocks 1 (M = 4.78, SD = 8.15) through 5 (M = 5.89, SD = 16.22) than in blocks 6 (M = 70.56, SD = 33.23) through 9 (M = 70.56, SD = 42.28). Finally, boredom increased as a function of time (F (1, 3.20) = 10.36, = 3477.77). A trend analysis on all the above variables indicates a linear trend.

In addition, an analysis was performed to determine whether time had an effect on the individual FIT parameters and the overall FIT index. This was done by plotting the data and looking at the appropriate responses expected and performing ANOVAs on measures of the FIT index, initial pupil diameter, saccadic velocity, latency of pupil constriction, and constriction amplitude of pupil. The FIT parameters that appeared to be affected by time were initial pupil diameter (F (1, 5.02) = 4.29, MSE = 4.05) and latency of pupil constriction (F (1, 7.58) = 2.88, MSE = 2.84). These data are represented graphically in Figures 7 and 8.

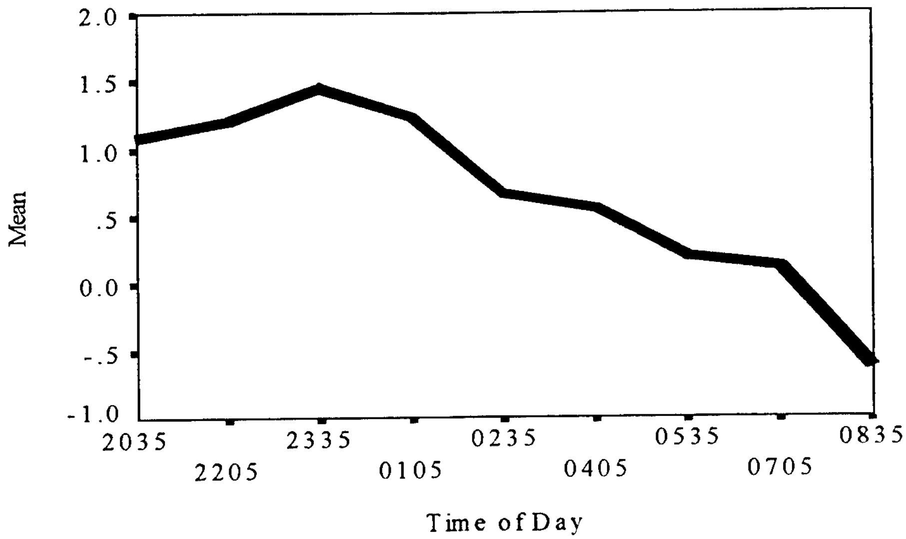

Figure 7. Initial pupil diameter across time of day.

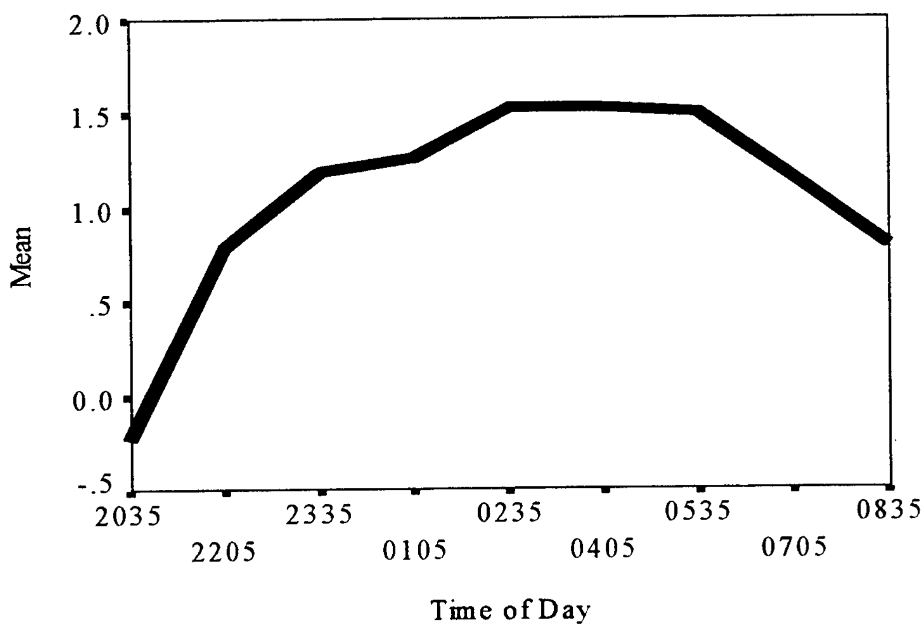

Figure 8. Latency of pupil constriction across time of day.

Overall, the manipulation check showed that performance and mood were affected by time. Lapses and RT increased and d' decreased as participants reported feeling more fatigued, sleepy, and bored, and less attentive and alert. Time did not seem to have an effect on the FIT index. However, two individual parameters that make up that index were affected by time. These were initial pupil diameter and latency of pupil constriction.

To establish criterion related validity, a correlational analysis of the FIT index, d', and lapses was conducted. Table 2 displays those correlations. There should be a negative correlation between the FIT index and d'. As the FIT index increases over time, d' should decrease. As can be seen from the table, none of the correlations proved significant. Correlations between the FIT index and lapses and FIT index with RT should be positive. As the FIT index increases over time, so should lapses and RT. Again, looking at Table 2, significant correlations were in the opposite direction.

Table 2. Correlation Coefficients between the FIT Index and the Dependent Measures Across Blocks									
	FIT Index Block 1	**FIT Index** Block 2	**FIT Index** Block 3	**FIT Index** Block 4	**FIT Index** Block 5	**FIT Index** Block 6	**FIT Index** Block 7	**FIT Index** Block 8	**FIT Index** Block 9
d'	-.52 p=.15	-.34 p=.36	-.28 p=.46	.06 p=.89	-.09 p=.82	.05 p=.89	-.25 p=.52	-.59 p=.09	.33 p=.39
lapses	.29 p=.44	.33 p=.38	.67 p=.50	-.44 p=.24	-.67 p=.05	-.51 p=.16	-.52 p=.15	.04 p=.93	-.51 p=.16
RT	-.69 p=.04	-.39 p=.29	-.55 p=.12	-.40 p=.29	-.59 p=.10	-.65 p=.06	-.58 p=.10	-.14 p=.72	-.55 p=.13

* Note: df=8 for all correlations.

Because the FIT index was not found to be a valid predictor of the dependent measures (lapses, d', RT), the next step was a correlational analysis of latency of pupil constriction and d', lapses, fatigue, boredom, sleepiness, attention, and alertness. For convergent validity, a correlational analysis was conducted of latency of pupil constriction versus fatigue, boredom, and sleepiness. For divergent validity, an analysis was conducted for latency of pupil constriction versus alertness and attention. In the following analyses, initial pupil diameter was not used even though it was statistically significant because participants began testing at one standard deviation above their baseline, as indicated in Figure 7. This would seem to indicate that there might have been slight impairment at the start of testing. There was a significant correlation at 0600 hours between fatigue and latency of pupil constriction $r(8) = .783$, $p < .013$ and between sleepiness and latency of pupil constriction ($r(8) = .753$, $p < .019$). No other correlations were significant.

Correlations computed between lapses, d', RT, and the subjective measures were on the whole not significant. However, d' and self-reported measures of attention were significantly related at 0430 hours ($r(8) = .672$, $p < .047$). The measures of d' and sleepiness were significantly related at 0300 hours ($r(8) = -.849$, $p < .004$) and at 0600 hours ($r(8) = .654$, $p < .05$). Also, at 0105 hours, RT and the self reported measures of sleepiness and fatigue were significantly related ($r(8) = .783$, $p < .013$ and $r(8) = .678$, $p < .04$).

DISCUSSION

As predicted, participants experienced decrements in performance and mood across time blocks and became increasingly fatigued. As the night progressed, participants became less accurate in responding to the stimuli and showed increases in numbers of lapses (non-response). Also, participants reported, on the VAS, that they experienced increased sleepiness and boredom and that fatigue was beginning to set in as the night progressed. Attention and alertness as reported on the VAS dramatically decreased across time. Thus, it can be concluded that participants experienced the negative effects of sleep loss, consistent with previous research indicating that sleep deprivation produces decreases in performance levels and changes in mood states (Akerstedt, 1991; Krueger, 1989; Mast & Heimstra, 1964; Moore-Ede, 1993; Stanny et al., 1993; Wiegmann et al., 1993).

It was anticipated that the FIT index and individual parameters would be affected by time; however, this was not the case. The FIT index performed sporadically and did not prove to be a valid measure of impairment due to fatigue. Saccadic velocity was expected to be a high early in the evening and, as the night progressed, it was expected to slow. Instead, saccadic velocity varied sporadically over time for the group as a whole. The same apparently random pattern over time was obtained for pupillary constriction.

As expected, initial pupil diameter and latency of pupil constriction were significantly affected by time. The literature indicates that, as fatigue sets in, the diameter of the pupil becomes smaller and the rate of constriction gets slower (Andreassi, 1989). In this study, the initial pupil diameter and latency of pupil constriction followed these patterns. However, for the remainder of the analysis, initial pupil diameter was not used because participants were observed to be above their baselines when tested on the first block of the evening. A possible explanation for this phenomenon is that pupil dilation might follow a circadian rhythm, as do many bodily functions. As day turns into night, physiological adjustments in the eye must be made in order to see properly. This means that the pupil must dilate to allow more light into the retina. It may be the case that as night falls, the pupil automatically increases in size. So, even if we are indoors with the lights on our pupils may be larger because of this natural function or circadian rhythm.

Latency of pupil constriction was positively correlated with fatigue and sleepiness as the literature suggests (Andreassi, 1989). As latency of pupil constriction increased, participants reported feeling more fatigued and sleepy. Also when the performance measures (d' and lapses) were compared with the subjective measures, d' was positively correlated with attention and was negatively correlated with sleepiness, as suggested by the literature (Akerstedt, 1991; Krueger, 1989; Mast & Heimstra, 1964; Moore-Ede, 1993; Stanny et al., 1993; Wiegmann et al., 1993).

In this study, participants experienced behavioral effects of sleep loss in only a 24-hour sleep deprivation period. These effects were manifested in both subjective measures and objective performance measures. The finding that the FIT data did not correlate with these measures was unexpected. However, the latency of pupil constriction and initial pupil diameter demonstrated promise as predictive indexes of fatigue. Future research may corroborate these findings.

Pulse Medical Instruments (1994) validated the FIT in field tests under a variety of conditions including sleep deprivation, fatigue, and influences of different types of drugs and alcohol. The sleep deprivation period used in one study reported consisted of a 72-hour or greater period. Differences in design may explain the differences between the results obtained in their validation studies and the present results. FIT may be capable of detecting impairment due to fatigue and sleep deprivation over long periods, but not be sensitive enough to detect impairment resulting from a 24-hour period of deprivation. From a practical standpoint, the present 24-hour sleep deprivation paradigm more closely resembles fatigue produced by actual shift work schedules. The present study has tried to more accurately measure the utility of FIT for Fitness-for Duty evaluations.

A valid criticism of the current study is the small sample size, which raises the question of power. However, sample size was adequate to obtain statistically significant results for both the performance measures and the subjective measures in the manipulation check, which tends to somewhat counter arguments based on inadequate power.

Shiftwork has been shown to disrupt circadian rhythms, which can lead to sleep loss, health problems, decreased performance, increased moodiness, and problems in safety (Angus et al., 1985; Krueger, 1989; Mast & Heimstra, 1964; Moore-Ede, 1993; Rosekind et al., 1994; Stanny et al., 1993; Wiegmann et al., 1993). Based on these findings, it is important to have some measure that can detect impairment due to fatigue. Current methods require either long testing periods, which are costly, or are measures of behavioral performance, in which an individual has the opportunity to intentionally set artificially low baselines. To remedy these problems, the author believes that physiological measurements may prove to be the best method to detect impairment in employees. At this point, the FIT may not be the best way to do this. The data from this study indicate that behavioral and self-report measures are very reliable measures of fatigue and/or impairment. These types of methods may be the best way to proceed until a more reliable physiological measure can be designed.

Footnotes

[1] There was a change in the original formula that was used by PMI. Early analysis done using the original formula produced equivalent results. The new formula converted the data to measures that were normally distributed so that correlation coefficients could be calculated. The correlation coefficient between measures computed using the original formula and the new formula is 1.00.

[2] Throughout the analysis, the Huynh-Feldt epsilon adjustment for the degrees of freedom was used for F. It is believed to be a better adjustment than the Greenhouse-Geisser adjustment, which can be overly conservative, especially for small sample sizes (SPSS Windows for Advanced Statistics, 1993). For a better explanation of the Huynh-Feldt adjustment, see Huynh-Feldt (1976).

REFERENCES

Akerstedt, T. (1991). Sleepiness at work: Effects of irregular work hours. In T.H. Monk (Ed.), *Sleep, sleepiness, and performance* (pp. 129-152). Chichester, England: Wiley.

Akerstedt, T., & Gillberg, M. (1982). Displacement of the sleep period and sleep deprivation: Implications for shiftwork. *Human Neurobiology*, 1, 163-171.

Andreassi, J.L. (1989). *Psychophysiology: Human behavior and physiological response* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.

Angus, R.G., Heslegrave, R.J. & Myles, W.S. (1985). Effects of prolonged sleep deprivation, with and without chronic physical exercise, and mood and performance. *Psychophysiology*, 22, 276-282.

Colligan, M.J. & Tepas, D.I. (1986). The stress of hours of work. *American Industrial Hygiene Association Journal*, 47, 686-695.

DeJohn, C.A., Marr, M.J., Molina, E.A., & McCardie, A.H. (1992). *A computer-based visual analog scale* (Tech. Memo No 92-1). Pensacola, FL: Naval Aerospace Medical Research Laboratory.

Gililand, K., & Schlegel, R.E. (1994, May). *Development of a laboratory model of readiness to perform testing.* Paper presented at the Department of Defense Human Factors Engineering Technical Group Meeting 32, Oklahoma City, OK.

Gold, D.R., Rogacz, S., Bock, N., Tosteson, T.D., Baum, T.M., Speizer, F.E., & Czeisler, C.A. (1992). Rotating shift work, sleep, and accidents related to sleepiness in hospital nurses. *American Journal of Public Health*, 82, 1011-1014.

Huynh, H., & Feldt, L.S. (1976). Estimation of the Box correction for degrees of freedom from sample data in randomized block and split-plot designs. *Journal of Educational Statistics*, 1, 69-82.

Kay, G.G. (1995). CogScreen-aeromedical edition: Professional manual. *Psychological Assessment Resources, Inc.*

Krueger, G.P. (1989). Sustained work, fatigue, sleep loss and performance: A review of the issues. *Work and Stress*, 3, 129-141.

Mast, T.M., & Heimstra, N.W. (1964) Effects of fatigue on vigilance performance. *Journal of Engineering Psychology*, 3, 73-79.

Moore-Ede, M. (1993). *The twenty four hour society*. (W. Patrick, Ed.). NY: Addison-Wesley Publishing Company.

Monk, T.H. (1991). Circadian aspects of subjective sleepiness: a behavioural messenger? In T.H. Monk (Ed.), *Sleep, sleepiness and performance* (pp. 39-63). Chichester, England: Wiley.

Monk, T.H., & Folkard, S. (1983). Circadian rhythms and shiftwork. In G.R.J. Hockey (Ed.), *Stress and Fatigue in Human Performance* (pp. 97-121). Chichester, England: Wiley.

Nicholson, A.N., & Stone, B.M. (1987). *Sleep and wakefulness: Handbook for flight medical officers* (2nd ed.). United Kingdom: Advisory Group for Aerospace Research and Development.

Performance Factors, Incorporated. (1993). FACTOR 1000 fitness for work assessment program technology overview and technical validation. Englewood, CO: Performance Factors, Inc.

Pulse Medical Instruments, Inc. (1994, December). *FIT Workplace safety screener: Theory and validation.* Rockville, MD.

Rosekind, M.R., Gander, P.H., Miller, D.L., Gregory, K.B., Smith, R.M., Weldon, K.J., Co, E.L., McNally, K.L., & Lebacqz, J.V. (1994). Fatigue in operational settings: Examples from the aviation environment. *Human Factors,* 36, 327-338.

SPSS for Windows: Advanced Statistics 6.0. (1993). Chicago, Il.

Stanny, R.R., McCardie, A.H., & Neri, D.F. (1993). *Effects of methamphetamine on vigilance and tracking during extended wakefulness* (Tech Memo NO 1387) Pensacola, Fl: Naval Aerospace Medical Research Laboratory.

Wiegmann, D.L., Neri, D.F., Stanny, R.R., Shappell, S.A., McCardie, A.H., & McKay, D.L. (1993). *Behavioral effects of tyrosine during sustained wakefulness* (Tech Memo NO 1388) Pensacola, Fl: Naval Aerospace Medical Research Laboratory.

OTHER NAMRL RELATED PUBLICATIONS

Cited above.

Appendix A

Visual Analog Scale Items

Fluttering Heart
Dizzy
Ill at ease
Euphoric (feeling that all is well)
Overstimulated
Restless
Shaky
Headache
Diarrhea
Constipated
Abdominal cramps
Dry mouth
Unpleasant taste
Itchy
Fatigued
Depressed
Anxious
Hostile
Jumpy
Talkative
Hungry
Difficult urinating
Urinating frequently
Sleepy
Rapid breathing
Alert
Able to focus attention
Bored

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY *(Leave blank)*

2. REPORT DATE
27 April 1998

3. REPORT TYPE AND DATES COVERED

4. TITLE AND SUBTITLE
The First Fitness Impairment Test (FIT):
A First Look

5. FUNDING NUMBERS
63706N M0096.002-7010
DN248502

6. AUTHOR(S)
C. A. Portman-Tiller

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Naval Aerospace Medical Research Laboratory
51 Hovey Road
Pensacola Fl 32508-1046

8. PERFORMING ORGANIZATION REPORT NUMBER
NAMRL-1401

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Naval Medical Research and Development Command
Building 1, Tower 12
8901 Wisconsin Avenue
Bethesda, MD 20889-5606

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited.

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*
In a technologically advanced society with increasing demands, shiftwork has become the norm. Although shiftwork has become an accepted way of life, it has disadvantages, such as increases in fatigue, depressed mood states, and decreases in performance. The object of the study was to validate a test developed to determine whether or not a worker is fit for duty. The Fitness Impairment Test (FIT) is based on ocular measures that are believed to be correlated with fatigue. Participants were 9 Navy and Marine Corps aviation candidates, ages 23 to 26 years. ANOVA and correlational analyses were used to compare subjective and objective measures of performance and fatigue with the measures produced by FIT. Although participants showed decreases in performance and mood states, few statistically significant correlations were found between the physiological and performance measures.

14. SUBJECT TERMS
Fitness Impairment Test; FIT; Circadian rhythm; Readiness to perform tests; Validation study; Shiftwork

15. NUMBER OF PAGES
19

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	SAR